



0590  
0124

## **RAW SEQUENCE LISTING** **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/862,855C  
Source: OIPE  
Date Processed by STIC: 1/24/03

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:**

**<http://www.uspto.gov/web/offices/pac/checker>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. **EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>) , EFS Submission User Manual - ePAVE)**
2. **U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202**
3. **Hand Carry directly to:**  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
**Or**  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202**

**Raw Sequence Listing Error Summary**

**ERROR DETECTED**      **SUGGESTED CORRECTION**      **SERIAL NUMBER:** 09/862,855C

**ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE**

- 1 Wrapped Nucleic  
Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino  
Numbering The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) \_\_\_\_\_ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0  
"bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) \_\_\_\_\_. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences  
(OLD RULES) Sequence(s) \_\_\_\_\_ missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences  
(NEW RULES) Sequence(s) \_\_\_\_\_ missing. If intentional, please insert the following lines for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 9 Use of n's or Xaa's  
(NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ✓ Invalid <213>  
Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) \_\_\_\_\_ missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0  
"bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



OIPE

**Does Not Comply**  
Corrected Claims Needed

**RAW SEQUENCE LISTING**

PATENT APPLICATION: US/09/862,855C

DATE: 01/24/2003

TIME: 13:22:32

Input Set : A:\94652 sequence listing.txt  
 Output Set: N:\CRF4\01242003\I862855C.raw

```

4 <110> APPLICANT: Cai, Hong
5   Keller, Richard
6   Werner, James
7   Goodwin, Peter
9 <120> TITLE OF INVENTION: RAPID HAPLOTYPEING BY SINGLE MOLECULE DETECTION
11 <130> FILE REFERENCE: S-94,652
13 <140> CURRENT APPLICATION NUMBER: US 09/862,855C
C--> 14 <141> CURRENT FILING DATE: 2003-01-13
16 <150> PRIOR APPLICATION NUMBER: US 60/206,512
17 <151> PRIOR FILING DATE: 2000-05-22
19 <160> NUMBER OF SEQ ID NOS: 21
21 <170> SOFTWARE: PatentIn version 3.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 20
25 <212> TYPE: DNA
26 <213> ORGANISM: M13mp18 target containing the EcoR I restriction site
28 <400> SEQUENCE: 1
29 gctcgaaattc gtaatcatcg
32 <210> SEQ ID NO: 2
33 <211> LENGTH: 18
34 <212> TYPE: DNA
35 <213> ORGANISM: M13mp18 target containing the Hind III restriction site
37 <400> SEQUENCE: 2
38 cagtgc当地 cttcgatg
41 <210> SEQ ID NO: 3
42 <211> LENGTH: 97
43 <212> TYPE: DNA
44 <213> ORGANISM: MLL (HRX, Htrx) and AF4 (FEL) gene fusion
46 <400> SEQUENCE: 3
47 gaagttccca aaaccactcc tagtgagccc aaaaaaaaagc agcctccacc accaaaaacaa
49 tatgatacat cttcaaaaac tcactcaa at tctcagc
52 <210> SEQ ID NO: 4
53 <211> LENGTH: 27
54 <212> TYPE: DNA
55 <213> ORGANISM: MLL 3968L20
57 <400> SEQUENCE: 4
58 aaaaaatttct tgggcttcac tagggag
61 <210> SEQ ID NO: 5
62 <211> LENGTH: 29
63 <212> TYPE: DNA
64 <213> ORGANISM: AF4 4025L24
66 <400> SEQUENCE: 5
67 aaaaaaattt gagtgatg ttgaagatg

```

20  
18  
60  
97  
27  
29

*Invalid response, see error  
summary sheet item 10*

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/862,855C

DATE: 01/24/2003

TIME: 13:22:32

Input Set : A:\94652 sequence listing.txt  
 Output Set: N:\CRF4\01242003\I862855C.raw

```

70 <210> SEQ ID NO: 6
71 <211> LENGTH: 12
72 <212> TYPE: DNA
73 <213> ORGANISM: MLLCy5P
75 <400> SEQUENCE: 6
76 tttcttgggc tc                                12
79 <210> SEQ ID NO: 7
80 <211> LENGTH: 12
81 <212> TYPE: DNA
82 <213> ORGANISM: AF4FAMP
84 <400> SEQUENCE: 7
85 tttagtgtgag tt                               12
88 <210> SEQ ID NO: 8
89 <211> LENGTH: 12
90 <212> TYPE: DNA
91 <213> ORGANISM: MLLCy5L
93 <400> SEQUENCE: 8
94 tttcttgggc tc                                12
97 <210> SEQ ID NO: 9
98 <211> LENGTH: 12
99 <212> TYPE: DNA
100 <213> ORGANISM: AF4RGXL
102 <400> SEQUENCE: 9
103 tttagtgtgag tt                               12
106 <210> SEQ ID NO: 10
107 <211> LENGTH: 32
108 <212> TYPE: DNA
109 <213> ORGANISM: A*02011/A/TT/GT
111 <400> SEQUENCE: 10
112 tggcagctca gaccaccaag cacaagtggg ag      32
115 <210> SEQ ID NO: 11
116 <211> LENGTH: 76
117 <212> TYPE: DNA
118 <213> ORGANISM: A*02011/A/TT/GT
120 <400> SEQUENCE: 11
121 gcggcccatg tggcgagca gttgagagcc tacctggagg gcacgtgcgt ggagtggctc 60
123 cgtagatacc tggaga                            76
126 <210> SEQ ID NO: 12
127 <211> LENGTH: 32
128 <212> TYPE: DNA
129 <213> ORGANISM: A*0212/A/CA/GT
131 <400> SEQUENCE: 12
132 tggcagctca gaccaccaag cacaagtggg ag      32
135 <210> SEQ ID NO: 13
136 <211> LENGTH: 76
137 <212> TYPE: DNA
138 <213> ORGANISM: A*0212/A/CA/GT
140 <400> SEQUENCE: 13
141 gcggcccatg tggcgagca gcagagagcc tacctggagg gcacgtgcgt ggagtggctc 60

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/862,855C

DATE: 01/24/2003

TIME: 13:22:32

Input Set : A:\94652 sequence listing.txt  
 Output Set: N:\CRF4\01242003\I862855C.raw

|     |  |    |
|-----|--|----|
| 143 | cgcagatacc tggaga  | 76 |
| 146 | <210> SEQ ID NO: 14  |    |
| 147 | <211> LENGTH: 32   |    |
| 148 | <212> TYPE: DNA  |    |
| 149 | <213> ORGANISM: A*0236/A/TT/CG                                   |    |
| 151 | <400> SEQUENCE: 14   |    |
| 152 | tggcagctca gaccaccaa gacaagtggg ag                               | 32 |
| 155 | <210> SEQ ID NO: 15  |    |
| 156 | <211> LENGTH: 76   |    |
| 157 | <212> TYPE: DNA  |    |
| 158 | <213> ORGANISM: A*0236/A/TT/CG                                   |    |
| 160 | <400> SEQUENCE: 15   |    |
| 161 | gcggcccatg tggcgagca gttgagagcc tacctggagg gcacgtgcgt ggacgggctc | 60 |
| 163 | cgcagatacc tggaga  | 76 |
| 166 | <210> SEQ ID NO: 16  |    |
| 167 | <211> LENGTH: 32   |    |
| 168 | <212> TYPE: DNA  |    |
| 169 | <213> ORGANISM: A*2402101/G/CA/CG                                |    |
| 171 | <400> SEQUENCE: 16   |    |
| 172 | tggcagctca gaccaccaa ggcaagtggg ag                               | 32 |
| 175 | <210> SEQ ID NO: 17  |    |
| 176 | <211> LENGTH: 76   |    |
| 177 | <212> TYPE: DNA  |    |
| 178 | <213> ORGANISM: A*2402101/G/CA/CG                                |    |
| 180 | <400> SEQUENCE: 17   |    |
| 181 | gcggcccatg tggcgagca gcagagagcc tacctggagg gcacgtgcgt ggacgggctc | 60 |
| 183 | cgcagatacc tggaga  | 76 |
| 186 | <210> SEQ ID NO: 18  |    |
| 187 | <211> LENGTH: 32   |    |
| 188 | <212> TYPE: DNA  |    |
| 189 | <213> ORGANISM: A*24031/G/CA/GT                                  |    |
| 191 | <400> SEQUENCE: 18   |    |
| 192 | tggcagctca gaccaccaa ggcaagtggg ag                               | 32 |
| 195 | <210> SEQ ID NO: 19  |    |
| 196 | <211> LENGTH: 76   |    |
| 197 | <212> TYPE: DNA  |    |
| 198 | <213> ORGANISM: A*24031/G/CA/GT                                  |    |
| 200 | <400> SEQUENCE: 19   |    |
| 201 | gcggcccatg tggcgagca gcagagagcc tacctggagg gcacgtgcgt ggagtggctc | 60 |
| 203 | cgcagatacc tggaga  | 76 |
| 206 | <210> SEQ ID NO: 20  |    |
| 207 | <211> LENGTH: 32   |    |
| 208 | <212> TYPE: DNA  |    |
| 209 | <213> ORGANISM: A*2413/G/TT/GT                                   |    |
| 211 | <400> SEQUENCE: 20   |    |
| 212 | tggcagctca gaccaccaa ggcaagtggg ag                               | 32 |
| 215 | <210> SEQ ID NO: 21  |    |
| 216 | <211> LENGTH: 76   |    |
| 217 | <212> TYPE: DNA  |    |

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/862,855C

DATE: 01/24/2003

TIME: 13:22:32

Input Set : A:\94652 sequence listing.txt  
Output Set: N:\CRF4\01242003\I862855C.raw

218 &lt;213&gt; ORGANISM: A\*2413/G/TT/GT

220 &lt;400&gt; SEQUENCE: 21

221 gcggcccatg tggcgagca gttgagagcc tacctggagg gcacgtgcgt ggacgggctc 60

223 cgcagataacc tggaga 76

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/862,855C

DATE: 01/24/2003

TIME: 13:22:33

Input Set : A:\94652 sequence listing.txt  
Output Set: N:\CRF4\01242003\I862855C.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date